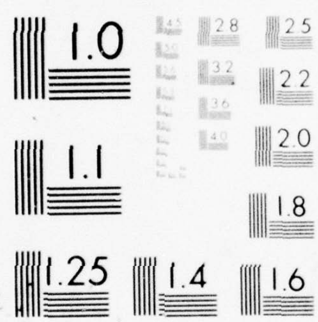
AD-A033 717 SYRACUSE UNIV N Y F/G 12/1
SOME FURTHER CONSIDERATIONS ON THE ESTIMATION OF ERROR OF MISCL--ETC(U)
OCT 76 K MEHROTRA F30602-75-C-0121

UNCLASSIFIED RADC-TR-76-322 NL

1 OF 1
AD
A033717

END
DATE
FILMED
2-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

24 [redacted]

RADC-TR-76-322
Technical Report
October 1976

SOME FURTHER CONSIDERATIONS ON THE ESTIMATION OF ERROR OF MISCLASSIFICATION BASED ON THE DESIGN SET

Syracuse University

**Approved for public release;
distribution unlimited.**

**ROME AIR DEVELOPMENT CENTER
AIR FORCE SYSTEMS COMMAND
GRIFFISS AIR FORCE BASE, NEW YORK 13441**

This report has been reviewed by the RADC Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

This report has been reviewed and approved for publication.

APPROVED: Haywood E. Webb, Jr.

HAYWOOD E. WEBB, Jr.
Project Engineer

APPROVED: Robert D. Krutz

ROBERT D. KRUTZ, Col, USAF
Chief, Information Sciences Division

FOR THE COMMANDER: John P. Huss

JOHN P. HUSS
Acting Chief, Plans Office

Do not return this copy. Retain or destroy.

MISSION
of
Rome Air Development Center

RADC plans and conducts research, exploratory and advanced development programs in command, control, and communications (C^3) activities, and in the C^3 areas of information sciences and intelligence. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.

AMERICAN REVOLUTION BICENTENNIAL 1776-1976

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 18 RADC-TR-76-322
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER: 9
4. TITLE (and Subtitle): 6 SOME FURTHER CONSIDERATIONS ON THE ESTIMATION OF ERROR OF MISCLASSIFICATION BASED ON THE DESIGN SET
5. TYPE OF REPORT & PERIOD COVERED: Technical Report, June 1975 - June 1976
6. PERFORMING ORG. REPORT NUMBER: N/A
7. AUTHOR(s): 10 Kishan Mehrotra
8. CONTRACT OR GRANT NUMBER(s): 15 F30602-75-C-0121
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Syracuse University, Syracuse NY 13210
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 920W, 95670016 — Query
11. CONTROLLING OFFICE NAME AND ADDRESS: Rome Air Development Center (ISCP), Griffiss AFB NY 13441
12. REPORT DATE: 11 October 1976
13. NUMBER OF PAGES: 46 — 12 52p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Same
15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer:
Haywood E. Webb, Jr. (ISCP)

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Pattern Recognition, Statistics, Error Estimate Misclassification, Sample Size, Mean, Variance, Dimensionality

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Estimates of the variance of the estimated probability of error for linear classifiers under normal distributions are calculated as related to dimensionality and the number of features.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 339 600

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

bpg

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

1. INTRODUCTION:

Use of Fisher's linear classifiers is very common in discriminant analysis. The linear classifiers are optimum when the underlying distributions are gaussian with common covariances. In other cases, they are not optimum, but their simplicity compensates for the loss in performance.

The probability of error is the very key quantity in pattern recognition and therefore considerable literature exists regarding its calculation and estimation. As discussed in Fukunaga (1972) there are two cases of interest as far as the estimation of the probability of error is concerned. First, estimation of the probability of error when the classifier is given and a sample of N observations is also given. This problem is considerably easy and as shown by Highleyman (1962) [see Fukunaga (1972, page 145-6)] an unbiased, minimum variance estimator is given by the ratio of number of misclassified observations to N [when a prior probabilities of classes are not available; if these prior probabilities are known, slightly better estimator can be obtained]. The above is a general result, i.e. applies to all possible classifiers. Second, when a sample consisting of N observations is available and the classifier has to be estimated, along with the probability of the error of misclassification. Obviously, the esimator of the probability of misclassification depends on the given class distributions and the classifiers to be used. In this paper we will consider the linear classifiers and the gaussian

distributions only.

At this point, the following, general result, of Hills (1966) is worth recalling. Suppose $\varepsilon(\Theta_1, \Theta_2)$ denotes the probability of error, Θ_1 are the parameters of the distribution used to design the Bayes classifier and Θ_2 are the parameters for the distributions used to test the performance of this Bayes classifier. If, for $\Theta_1 = \Theta_2$, $\hat{\Theta}_N$ denotes the estimator of the parameters based on a sample of size N and if

$$E[\varepsilon(\Theta, \hat{\Theta}_N)] = \varepsilon(\Theta, \Theta)$$

then

$$E[\varepsilon(\hat{\Theta}, \hat{\Theta})] \leq \varepsilon(\Theta, \Theta) \quad .$$

Thus, if the sample is used to obtain the classifier and the estimator of the probability of misclassification, then the method provides an *optimistic estimator* i.e. the expected value of the estimator is smaller than the true value. Following Fukunaga (1972) we call this method the C-method. Thus, in the C-method a given sample is first used to obtain the classifier and then is used for testing its performance.

In another approach the given sample can be used to obtain the classifier and a fresh sample to test its performance. Among many possibilities available to us, which use this approach, the leaving-one-out method [Lachenbruch (1965)] is rather economical. In this method, the sample of N observation is divided in two parts consisting of (N-1) and 1 observations respectively. The first set is used to construct the classifier and the remaining observation is used for testing. The method is repeated N times. Throughout

this paper, we will denote this method by the U-method.

Fukunaga and Kessel (1971) showed that for any random sample from the gaussian distributions, if an observation of the sample is misclassified by the C-method then, it will be misclsssified by the U-method, but the converse need not be true. Thus, for any sample the estimate of the error probability will be smaller when the C-method is used compared to the U-method.

In this paper we consider the C-method and the U-method of estimation of the probability of error for the linear classifiers and the gaussian distributions. Our aim is to evaluate the mean square errors of estimators for the purpose of comparisons. Let us remind ourselves that for the purpose of comparisons of two estimators mean square error is a good measure. Until recently these estimators were compared, mostly, using their expectations. Although there are several numerical studies, thoretical results are only recent. See Sorem (1971, 1972), Dasgupta (1974). The results of Moore, Whitsitt and Landgrebe (1976) are also of interest to a realted problem. Fukunaga's (1972, page 159) empirical study is of immediate interest to us.

In Section 2 some basic notations are introduced. In Section 3 we present Foley's (1972) results regarding the expectation of the estimator of the probability of error for the C-method when the covariances are known and the results of John (1961) which are applicable to the U-method. In Section 4, the C-method and the U-method are considered when the common covariances are unknown and the expected values of the estimators are considered. In

Section 5 we consider the variance of the estimator for the error probability by the C-method and finally, in Section 6 the same is evaluated for the U-method. In both Sections 5 and 6 the covariances are assumed known. Computer programs to evaluate these variances are presented in the Appendix.

2. NOTATIONS AND PRELIMINARIES:

We consider the two class pattern recognition problem. These two classes are denoted by $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively. The corresponding p-variate gaussian densities are denoted by $\phi(\mu_1,\Sigma)$ and $\phi(\mu_2,\Sigma)$ respectively. For the sake of convenience we assume that each class has equal a-priori probability. A random sample of size m from the first class is denoted by $X_1,\ldots,X_m$ and similarly another, independent random sample from $\mathcal{C}_2$ by $Y_1,\ldots,Y_n$.

An observation X is classified as belonging to the class $\mathcal{C}_1$ if

$$(2.1) \qquad X'\Sigma^{-1}(\bar{X}-\bar{Y}) - \tfrac{1}{2}(\bar{X}+\bar{Y})'\,\Sigma^{-1}(\bar{X}-\bar{Y}) \geq 0$$

when only μ_1 and μ_2 are unknown, Σ is assumed known, $\bar{X}$ and $\bar{Y}$ denote the sample means, i.e.

$$(2.2) \qquad \bar{X} = m^{-1}\sum_{i=1}^{m} X_i,\ \bar{Y} = n^{-1}\sum_{j=1}^{n} Y_j .$$

If Σ is also unknown, then the above, linear classifier, has to be modified and X is classified to $\mathcal{C}_1$ provided

$$(2.3) \qquad X'\,S^{-1}(\bar{X}-\bar{Y}) - \tfrac{1}{2}(\bar{X}+\bar{Y})'\,S^{-1}(\bar{X}-\bar{Y}) \geq 0$$

where $\bar{X}$ and $\bar{Y}$ are defined by (2.2) and

$$(2.4) \qquad S = (m+n-2)^{-1}\left\{\sum_{i=1}^{m}(X_i-\bar{X})(X_i-\bar{X})' + \sum_{j=1}^{m}(Y_j-\bar{Y})(Y_j-\bar{Y})'\right\} .$$

Probabilities of misclassification are given by

$$(2.5) \qquad \rho_{ij} = P[X\epsilon\mathcal{C}_j \mid X\epsilon\mathcal{C}_i],\ i \neq j,\ i,j = 1,2 .$$

Our aim is to study the estimates of ρ_{ij}'s using classifiers defined in (2.1) and (2.3). We recall, at this time, that the classifiers

are themselves random. Due to symmetry of the problem, it is sufficient to consider estimates of only one of the ρ_{12} or ρ_{21}.

3. EXPECTED VALUES OF THE ESTIMATES OF THE PROBABILITY OF ERROR, Σ KNOWN:

In this section the C-method and the U-method are used to obtain estimates of the probability of error when Σ, the common covariance matrix is known. The expected values are obtained for these estimators. It can be easily seen that when Σ is assumed known, we can generalize these results to the case when the two classes are allowed to different covariance matrices. Moreover, without loss of generality, we can assume $\Sigma = I$. Thus, in the following sections the common covariance is taken to be the identity matrix.

The results of Section 3.1 were obtained by Foley (1972), our treatment is only slightly different. Section 3.2 contains results obtained by John (1961).

3.1 The C-method.

Let

$$T_i = \begin{cases} 1 & \text{if } X_i \text{ is classified as an element of class } 2 \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

when the classifier (2.1) is employed. Then, an estimate of ρ_{21} is given by

$$\varepsilon_1 = m^{-1} \sum_{i=1}^{m} T_i \,. \tag{3.2}$$

Clearly, by symmetry in X's

$$(3.3)\quad \begin{aligned} E(\varepsilon_1) &= E(T_1) = P[(X_1-\tfrac{1}{2}(\bar{X}+\bar{Y}))'(\bar{X}-\bar{Y}) < 0] \\ &= E_{\bar{X},\bar{Y}}[P\{(X_1-\tfrac{1}{2}(\bar{X}+\bar{Y}))'(\bar{X}-\bar{Y}) < 0|\bar{X},\bar{Y}\}]. \end{aligned}$$

But, can be easily verified that the conditional distribution of X_1, given $\bar{X},\bar{Y}$ is gaussian with mean vector $\bar{X}$ and covariance matrix $\{(m-1)/m\}I$. Thus, the conditional distribution of $(X_1-\frac{1}{2}(\bar{X}+\bar{Y}))'(\bar{X}-\bar{Y})$, given $\bar{X},\bar{Y}$ will also be univariate gaussian with mean $1/2(\bar{X}-\bar{Y})'(\bar{X}-\bar{Y})$ and variance $((m-1)/m)(\bar{X}-\bar{Y})'(\bar{X}-\bar{Y})$. But, $Z = (\bar{X}-\bar{Y})'(\bar{X}-\bar{Y})$ is itself a random variable satisfying the noncentral chisquare distribution with p degrees of freedom and noncentrality parameter

$$\lambda^2 = \frac{mn}{m+n}(\mu_1-\mu_2)'(\mu_1-\mu_2) = \frac{mn}{m+n}\delta^2$$

where δ^2 is the usual Mahalnobis distance between the two gaussian distributions. Let $g(z;\lambda)$ denotes the density of Z, i.e.,

$$g(z;\lambda) = \sum_{k=0}^{\infty} e^{-(\frac{\lambda^2}{2})}\left(\frac{\lambda^2}{2}\right)^k \frac{1}{k!}\,\frac{1}{\Gamma(\frac{p+2k}{2})\,2^{\frac{p+2k}{2}}}\, e^{-\frac{z}{2}}\, z^{\frac{p+2k}{2}-1}.$$

Thus, from (3.3) and the above discussion,

$$(3.4)\quad E(\varepsilon_1) = \int_0^{\infty}\left[\int_{-\infty}^{0}\phi(x;\tfrac{1}{2}z,\tfrac{m-1}{m}z)\,dx\right]g(z,\lambda)\,dz$$

where $\phi(x;\mu,\sigma^2)$ denotes the univariate gaussian density with mean μ and variance σ^2. At this stage the order of integration and summation can be interchanged and also the double integration can be evaluated after changing to polar coordinate system. The final result is obtained in the form of an infinite series given below.

(3.5) $$E(\varepsilon_1) = \sum_{k=0}^{\infty} \exp[-\frac{\lambda^2}{2}](\frac{\lambda^2}{2})^k \frac{\Gamma(k+\frac{p}{2}+1)}{\Gamma(\frac{1}{2})\,\Gamma(k+\frac{p}{2})} I(p+2k,m+n)$$

where

$$I(M,N) = \int_0^A \sin^{M-1}\theta \, d\theta$$

$$A = \tan^{-1}\sqrt{2(N-1)} \quad .$$

3.2 <u>The U-method</u>:

In this case, we first obtain the linear classifier using the (N-1) observations and then test the classifier on the remaining observation. The process is repeated N times. An estimate of ρ_{21} will be given by

(3.6) $$\varepsilon_2 = \sum_{i=1}^{m} T_i^*/m$$

where

(3.7) $$T_i^* = \begin{cases} 1 & \text{if } X_i \text{ is misclassified to class } 2 \\ 0 & \text{otherwise.} \end{cases}$$

where classifier (2.1) is used after replacing $\bar{X}$ by $\bar{X}_{(i)}$ and X equals X_i; $\bar{X}_{(i)} = (m-1)\ [m\bar{X}-X_i]$. Once again, due to symmetry of the problem

$$E(\varepsilon_2) = E(T_1^*) = P[(X_1-\frac{1}{2}(\bar{X}_{(i)}+\bar{Y}))'(\bar{X}_{(i)}-\bar{Y}) < 0] \ .$$

However, the above is the same probability which was obtained by John (1961). The similarity follows, as soon as we recognize the fact that X_i, $\bar{X}_{(i)}$ and $\bar{Y}$ are all independently distributed. Thus, from equation (77) of John (1961) we get

$$(3.8)\quad E(\varepsilon_2) = e^{-(\lambda_1+\lambda_2)} \left[\sum_{r=0}^{\infty} \sum_{s=0}^{r} \frac{\lambda_1^r \lambda_2^s}{r!s!} \{1 - I_{\frac{1}{2}(1-\rho)}(\tfrac{1}{2}p+r, \tfrac{1}{2}p+s)\} \right.$$

$$\left. + \sum_{r=0}^{\infty} \sum_{s=r+1}^{\infty} \frac{\lambda_1^r \lambda_2^s}{r!s!} I_{\frac{1}{2}(1+\rho)}(\tfrac{1}{2}p+s\ \tfrac{1}{2}p+r) \right],$$

where

$$\lambda_1 = \frac{\delta^2(m-1)n}{4(1+\rho)} \left[\frac{1}{(m+n-1)^{\frac{1}{2}}} - \frac{1}{\{m+n-1+4(m-1)n\}^{\frac{1}{2}}} \right]^2,$$

$$\lambda_2 = \frac{\delta^2(m-1)n}{4(1-\rho)} \left[\frac{1}{(m+n-1)^{\frac{1}{2}}} + \frac{}{\{m+n-1+4(m-1)n\}^{\frac{1}{2}}} \right]^2,$$

$$\rho = \frac{m-n-1}{(m+n-1)\{m+n-1+4(m-1)n\}^{\frac{1}{2}}}$$

and

δ^2 is, as defined above, the Mahalnobis distance.

4. EXPECTED VALUE OF THE ESTIMATE OF THE PROBABILITY OF ERROR: Σ UNKNOWN:

In this section the results of the previous section are extended for the case when Σ is unknown. In section 4.1 the C-method is considered and in 4.2 the U-method. In the section 4.2 it is shown that no new result is needed since results of Okomoto (1963) apply. Results of the section 4.1 are new.

4.1 The C-method:

Since Σ is assumed to be unknown, it must be estimated and S defined in (2.4) provides an estimate. An estimate of ρ_{21} is clearly given by

$$\varepsilon_3 = m^{-1} \sum_{i=1}^{m} V_i \tag{4.1}$$

where

$$V_i = \begin{cases} 1 & \text{if } \{X_i - \frac{1}{2}(\bar{X}+\bar{Y})\}' S^{-1}(\bar{X}-\bar{Y}) < 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

$i = 1,2,\ldots,m$. In this section we obtain an expression for the expected value of ε_3. The expression is then evaluated by means of numerical integration.

A conditional argument is employed to simplify the expected value of ε_3. Since $X_1,\ldots,X_m$ are independent and identically distributed

$$E(\varepsilon_3) = m^{-1} \sum_{i=1}^{m} E(V_i) = E(V_1)$$

$$= P[\{X_1 - \tfrac{1}{2}(\bar{X}+\bar{Y})\}' S^{-1}(\bar{X}-\bar{Y}) < 0]$$

$$= 1 - P[X_1' S^{-1}(\bar{X}-\bar{Y}) > \tfrac{1}{2}(\bar{X}+\bar{Y})' S^{-1}(\bar{X}-\bar{Y})]$$

$$(4.3) \qquad = 1 - E_{\bar{X},\bar{Y},S}[P\{X_1' S^{-1}(\bar{X}-\bar{Y}) > \tfrac{1}{2}(\bar{X}+\bar{Y})' S^{-1}(\bar{X}-\bar{Y}) \mid \bar{X},\bar{Y},S\}] .$$

Firstly we consider the inner conditional probability in (4.3) and secondly we obtain its expected value with respect to $\bar{X}$, $\bar{Y}$, S.

Results

Theorem 4.1. Let $U_1 = X_1' S^{-1} (\bar{X}-\bar{Y})$.

(i) The conditional distribution of U_1 given $\bar{X}$, $\bar{Y}$, S is

$$(4.4) \qquad f_4(u_1|\bar{X},\bar{Y},S) = \begin{cases} d\left(\frac{m-1}{m}B_{11}\right)^{-\frac{1}{2}} \left\{1 - \frac{m}{m-1}\frac{(u_1-\bar{u}_1)^2}{B_{11}}\right\}^{\frac{m+n-5}{2}} \\ \qquad \text{for } \frac{m}{m-1}\frac{(u_1-\bar{u}_1)^2}{B_{11}} \le 1 \\ 0 \quad \text{otherwise} \end{cases}$$

where d, $\bar{u}_1$ and B_{11} are given by (4.12) below.

$$(4.5) \quad \text{(ii)} \quad E(V_1) = \frac{1}{2} - \frac{1}{2} P[W \ge k^{-2}] - \sum_{j=0}^{\infty} C(j) \int_0^{k^{-2}} \int_0^{k\sqrt{w}} \{(1-t^2)^{\frac{m+n-5}{2}} w^{j+\frac{p}{2}-1} (1+w)^{-(j+\frac{m+n-1}{2})}\} dt\, dw$$

where k is given by (4.16), C(j) by (4.18) and $W = \chi_1^2/\chi_2^2$, where χ_1^2 is a noncentral chisquare random variable with p degrees of freedom and noncentrality parameter

$$\lambda_1^2 = \frac{mn}{m+n} (\mu_1-\mu_2)' \Sigma^{-1} (\mu_1-\mu_2),$$

and χ_2^2 is a chisquare random variable with (m+n-p-1) degrees of freedom.

Proof: (i) Set $A = (m+n-2)S$, $\overset{o}{X} = (m-1)^{-1} \sum_{i=2}^{m} X_i$

$$\overset{o}{A} = \sum_{i=2}^{m} (X_i-\overset{o}{X})(X_i-\overset{o}{X})' + \sum_{j=1}^{n} (Y_i-\bar{Y})(Y_i-\bar{Y})' .$$

Then, X_i, $\overset{o}{X}$ and $\overset{o}{A}$ are statistically *independently* distributed and the density of X_1 is $\phi(\mu_1,\Sigma)$, of $\overset{o}{X}$ is $\phi(\mu_1,(m-1)^{-1}\Sigma)$ and of $\overset{o}{A}$ is Wishart $(m+n-3,\Sigma)$. Consequently, the joint density of X_1, $\overset{o}{X}$ and $\overset{o}{A}$, denoted by f_1, is

$$(4.6) \quad f_1(x_1,\overset{o}{x},\overset{o}{A}) = c_1\{\exp-\tfrac{1}{2}(x-\mu_1)'\Sigma^{-1}(x-\mu_1)\}\{\exp-\tfrac{m-1}{2}(\overset{o}{x}-\mu_1)'\Sigma^{-1}(x-\mu_1)\}$$

$$\{|\overset{o}{A}|^{\frac{m+n-4-p}{2}} \exp -\tfrac{1}{2} \operatorname{trace} \Sigma^{-1} \overset{o}{A}\}$$

where the constant c_1 is given by

$$(4.7) \quad c_1 = \{(2\Pi)^{-\frac{p}{2}}|\Sigma|^{-\frac{1}{2}}\}\{(2\Pi)^{-\frac{p}{2}}|(m-1)^{-1}\Sigma|^{-\frac{1}{2}}\}\{2^{\frac{m+n-3}{2}p}$$

$$\pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p} \Gamma(\tfrac{m+n-2-i}{2})\}^{-1} |\Sigma|^{-\frac{m+n-3}{2}} .$$

Since, X, $\overset{o}{X}$, A and $\overset{o}{A}$ are related by the following equalities

$$\overset{o}{X} = (m-1)^{-1} (m\bar{X}-X_1)$$

$$\overset{o}{A} = A - m(m-1)^{-1} (X_1-\bar{X})(X_1-\bar{X})'$$

the joint density of X_1, $\bar{X}$ and A can be easily obtained from (4.7) by standard procedures. Moreover, we also know that the random variables $\bar{X}$ and A are statistically independent and that $\bar{X}$ follows a gaussian density $\phi(\mu_1, m^{-1}\Sigma)$ and the random variable A is distributed as Wishart $(m+n-2, \Sigma)$. Thus, the conditional density of X_1 given $\bar{X}$ and A is obtained by taking the ratio of joint density of X_1, $\bar{X}$ and A and of $\bar{X}$, A. This conditional density simplifies to:

$$(4.8) \quad f_2(x_1|\bar{x},A) = \begin{cases} c_2|A|^{-\frac{1}{2}}\{1-Q_1(x_1)\}^{\frac{m+n-p-4}{2}} & \text{for } Q_1(x_1) \le 1 \\ 0 & \text{otherwise} \end{cases}$$

where

$$(4.9) \qquad c_2 = \pi^{-\frac{p}{2}}\{m(m-1)^{-1}\}^{\frac{p}{2}}\,\Gamma\left(\frac{m+n-2}{2}\right)\left\{\Gamma\left(\frac{m+n-2-p}{2}\right)\right\}^{-1}$$

$$Q_1(x_1) = m(m-1)(x_1-\bar{x})'\,A^{-1}(x_1-\bar{x}) \quad .$$

In order to obtain the conditional density of U_1 given $\bar{X}$, $\bar{Y}$ and S, we first take any nonsingular square matrix T whose first row is given by $(\bar{X}-\bar{Y})'S^{-1}$, and make the transformation

$$U = T\,X_1 \quad .$$

Also denote TAT' by B and $T\bar{X}$ by $\bar{U}$. Then, clearly U_1 is the first element of the vector random variable U. From (4.8) the conditional distribution of U is easily obtained and is given by

$$
(4.10)\quad f_3(U|\bar{x},\bar{y},A) = \begin{cases} c_2|B|^{-\frac{1}{2}}\{1-Q_2(u)\}^{\frac{m+n-p-4}{}} & \text{for } Q_2(u) \le 1 \\ 0 & \text{otherwise,} \end{cases}
$$

where c_2 is given by (4.9) and

$$
Q_2(u) = m(m-1)^{-1}\ (u-\bar{u})'\ B^{-1}\ (u-\bar{u})\ .
$$

In order to obtain the conditional density of U_1 it remains to integrate out the last p-1 components of U. To perform this integration, we first partition,

$$
(4.11)\quad U = \begin{bmatrix} U_1 \\ \hdashline U^* \end{bmatrix},\quad \bar{U} = \begin{bmatrix} \bar{U}_1 \\ \hdashline \bar{U}^* \end{bmatrix},\quad B = \left[\begin{array}{c:c} B_{11} & B_{12} \\ \hdashline B_{21} & B_{22} \end{array}\right].
$$

Then,

$$
Q_2(u) = \frac{(u_1-\bar{u}_1)^2}{B_{11}} + \{(u^*-\bar{u}^*) - B_{21}B_{11}^{-1}\ (u_1-\bar{u}_1)\}'\{B_{22}-B_{21}B_{11}^{-1}B_{12}\}^{-1}
$$

$$
\{(u^*-\bar{u}^*) - B_{21}B_{11}^{-1}\ (u_1-\bar{u}_1)\},
$$

$$
= Q_3(u_1) + Q_4(u_1,u^*) \qquad \text{(say)}
$$

$$
|B| = B_{11}|B_{22} - B_{21}B_{11}^{-1}B_{12}|
$$

and

$$
Q_2(u) \le 1 \text{ if and if } Q_4(u_1,u^*) \le 1 - Q_3(u_1)\ .
$$

Thus, the integration of f_3, with respect to the vector u* can be performed easily, using the identity

$$\int_{\{t:(t't\leq 1)\}} (1-t't)^{\frac{M-p}{2}-1} dt = \pi^{\frac{p}{2}} \Gamma(\tfrac{M-p}{2})\{\Gamma(\tfrac{M}{2})\}^{-1}$$

and the linear transformation

$$t = [\{1-Q_3(u_1)\}(B_{22}-B_{21}B_{11}^{-1}B_{12})]^{-\frac{1}{2}} \{(u^*-\bar{u}^*) - B_{21}B_{11}^{-1}(u_1-\bar{u}_1)\} .$$

Therefore, the conditional density of U_1 given $\bar{X}$, $\bar{Y}$ and S'

is given by

$$f_4(u_1 \; \bar{X},\bar{Y},S) = \begin{cases} dB_{11}^{-1/2}\left\{1 - \dfrac{m}{m-1} \dfrac{(u_1-\bar{u}_1)^2}{B_{11}}\right\}^{\frac{m+n-5}{2}} & \text{for } \dfrac{m}{m-1}\dfrac{(u_1-\bar{u}_1)^2}{} \leq 1 \\ 0 & \text{otherwise,} \end{cases}$$

where

$$(4.12) \quad d = \frac{1}{\sqrt{\pi}}\Gamma(\tfrac{m+n-2}{2}) \{\Gamma(\tfrac{m+n-3}{2})\}^{-1}$$

B_{11} = (1,1)th element of $B = (m+n-2)(\bar{X}-\bar{Y})' \, S^{-1}(\bar{X}-\bar{Y})$

$\bar{U}_1 = \bar{X}' \, S^{-1} \; (\bar{X}-\bar{Y})$.

This completes the proof of (i).

(ii) By equation (4.3)

$$(4.13) \quad E(V_1) = 1 - E_{\bar{X},\bar{Y},S}[P\{(u_1-\bar{u}_1) > -\frac{B_{11}}{2(m+n-2)} | \; \bar{X},\bar{Y},S\}]$$

$$= 1 - E_{B_{11}}[P\{\sqrt{\frac{m}{(m-1)B_{11}}}\,(U_1-\bar{U}_1) > -\frac{1}{2(m+n-2)}\sqrt{\frac{mB_{11}}{(m-1)}} | B_{11}\}]$$

because the conditional density of U_1 depends on $\bar{X}$, $\bar{Y}$ and S only through B_{11}. To this end, we observe that B_{11} is a multiple of Hotelling's T^2 statistic [see Anderson , page 108] and the distribution of Hotelling's T^2 is well known to be the same as the distribution of the ratio of two chisquares. That is, the distribution of

$$W = \frac{mn}{(m+n)(m+n-2)^2} B_{11} \tag{4.14}$$

is the same as of χ_1^2/χ_2^2 where χ_1^2 is a noncentral chisquare random variable with p degrees of freedom and noncentrality parameter

$$\lambda_1^2 = \frac{mn}{m+n} (\mu_1-\mu_2)' \Sigma^{-1} (\mu_1-\mu_2) \tag{4.15}$$

and χ_2^2 is, another, central chisquare, independent of χ_1^2, and with (m+n-p-1) degrees of freedom. Thus by part (i) of the theorem, and by symmetry of the distribution of $\sqrt{m(m-1)^{-1}B_{11}^{-1}}\,(u_1-\bar{u}_1)$ about the origin, we get

$$P\left\{ \sqrt{\frac{m}{(m-1)B_{11}}}\,(u_1-\bar{u}_1) > -\frac{1}{2(m+n-2)} \sqrt{\frac{mB_{11}}{(m-1)}} \,\middle|\, B_{11} \right\}$$

$$= \begin{cases} \frac{1}{2} + d \int_0^{k\sqrt{w}} (1-t^2)^{\frac{m+n-5}{2}} dt & \text{if } k\sqrt{w} < 1 \\ 1 & \text{if } k\sqrt{w} \geq 1 \end{cases}$$

where

$$k = [(m+n)\{4n(m-1)\}^{-1}]^{1/2} \tag{4.16}$$

Therefore,

$$EV_1 = 1 - P[w \geq k^{-2}] - \frac{1}{2} P(W \leq k^{-2}) - [d \int_0^{k^{-2}} \int_0^{k\sqrt{w}} (1-t^2)^{\frac{m+n-5}{2}}$$

$$\{e^{-\frac{\lambda_1}{2}} \sum_{j=0}^{\infty} (\frac{\lambda_1}{2})^j \frac{1}{j!} \Gamma(\frac{m+n-1}{2}) \{\Gamma(\frac{m+n-p-1}{2}) \Gamma(j+\frac{p}{2})\}^{-1}$$

$$\frac{w^{\frac{p}{2}+j-1}}{(1+w)^{j+\frac{m+n-1}{2}}}\} dt\, dw]$$

$$(4.17) = \frac{1}{2} - \frac{1}{2}p(w \geq k^{-2}) - \sum_0^{\infty} c(j) \int_0^{k^{-2}} \int_0^{k\sqrt{w}} (1-t^2)^{\frac{m+n-5}{2}}$$

$$w^{\frac{p}{2}+j-1} (1+w)^{-\{\frac{m+n-1}{2}+j\}} dt\, dw$$

where the constant $c(j)$ depends on λ_1, j, m, n and p and is given by

$$c(j) = d\{e^{-\frac{\lambda_1}{2}} (\frac{\lambda_1}{2})^j \frac{1}{j!}\} \Gamma(\frac{m+n-1}{2}) \{\Gamma(\frac{m+n-p-1}{2}) \Gamma(j+\frac{p}{2})\}^{-1} . \tag{4.18}$$

Thus, the theorem is proved.

4.1.1 <u>Numerical Evaluation of $E(\varepsilon_3)$</u>

In its present form, Theorem 4.1, (ii) is not convenient to evaluate. In this sub-section we will obtain further simplifications. We will consider the simple case of $\delta = 0$ in detail. For $\delta = 0$, all of the terms of the infinite series of integrals are zero except

the first term. For $\delta \neq 0$, there will be infinite terms, however due to the coefficients $c(j)$, the series will be a fast converging series. Moreover, each term of the infinite series of integrals can be evaluated exactly in the same manner as the terms for $\delta = 0$. Thus, for $\delta \neq 0$ details are omitted and only the numerical values are given.

For $\delta = 0$, (4.17) gives

$$(4.19) \quad EV_1 = \frac{1}{2} - \frac{1}{2}P(w \geq k^{-2}) - c(0) \int_0^{k^{-2}} \int_0^{k\sqrt{w}} (1-t^2)^{\frac{m+n-5}{2}} w^{\frac{p}{2}-1} (1+w)^{-\frac{m+n-1}{2}} \, dt \, dw \, .$$

Since $P(w \geq k^{-2})$ can be obtained from the incomplete beta tables [reference [13]], we consider the integral involved in the third term. We will reduce this double integral into a sum of single integrals and the later are evaluated by numerical integrations. At this stage, we also make additional simplifying assumption m=n. As is clear from the following development that the general case $m \neq n$ can be handled exactly in the same manner. For m=n, the integral in the third term on the right hand side of equation (4.19) is given by

$$(4.20) \quad c(0) \int_0^{k^{-2}} \int_0^{k\sqrt{w}} (1-t^2)^{\frac{2n-5}{2}} w^{\frac{p}{2}-1} (1+w)^{-\frac{2n-1}{2}} \, dt \, dw.$$

We make the transformations $t^2 = u$ and $\frac{w}{1+w} = y$ so that (4.20) is equal to

$$\frac{c(0)}{2} \int_0^{(1+k^2)^{-1}} \int_0^{k^2 y(1-y)^{-1}} (1-u)^{\frac{2n-5}{2}} u^{-1/2} y^{\frac{p}{2}} (1-y)^{\frac{2n-p-3}{2}} \, du \, dy.$$

By changing the order of integration the above expression becomes

$$\frac{c(0)}{2} \int_0^1 \int_{u(1+k^2)^{-1}}^{(1+k^2)^{-1}} y^{\frac{p}{2}-1} (1-y)^{\frac{2n-p-3}{2}} (1-u)^{\frac{2n-5}{2}} u^{-1/2} \, dy \, du.$$

If p is odd then, (2n-p-3) is even and integration by parts can be performed to evaluate the above inner integral in a recursive manner. In the alternative case, i.e. when p is even, p/2-1 is an integer and therefore once again we can integrate by parts. Thus, (4.20) will be equal to

$$
(4.21)\quad
\begin{cases}
\dfrac{c(0)}{2} \displaystyle\sum_{i=0}^{a} \dfrac{a!}{(a-i)!\{(b+1)\dots(b+1+i)\}(1+k^2)^{a+b+1}} \\[1ex]
\quad \left\{\displaystyle\int_0^1 u^{a-i-1/2} (1-u)^{\frac{2n-5}{2}} (1+k^2-u)^{b+1-i}\, du - \dfrac{k^{2(b+1+i)}\,\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{2n-4}{2}\right)}{\left(\frac{2n-3}{2}\right)}\right\} \\[1ex]
\qquad \text{when } p \text{ is even; } a = \dfrac{p}{2} - 1,\ b = \dfrac{2n-p-3}{2} \\[2ex]
\dfrac{c(0)}{2} \displaystyle\sum_{i=0}^{b} \dfrac{b!}{(b-i)!(a+1)\dots(a+1+i)(1+k^2)^{a+b+1}} \\[1ex]
\quad \left\{\dfrac{k^{2(b-i)}\,\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{2n-4}{2}\right)}{\Gamma\left(\frac{2n-3}{2}\right)} - \displaystyle\int_0^1 u^{a+1+i-\frac{1}{2}} (1-u)^{\frac{2n-5}{2}} (1+k^2-u)^{b-i}\, du\right\} \\[1ex]
\qquad \text{when } p \text{ is odd, } a = \dfrac{p}{2} - 1,\ b = \dfrac{2n-p-3}{2}.
\end{cases}
$$

We evaluate the integrals involved in (4.21) by numerical integration. Some numerical values are presented below in Tables 4.1, 4.2 and 4.3.

Table 4.1

p \ N/p	1	2	3	4	5
3	1.00	.711	.663	.638	.629
5	.881	.713	.667	.642	.625
7	.852	.715	.668	.640	.624

Table 4.2

$\delta=1$

p \ N/p	1	2	3	4	5
3	1.0	.739	.687	.653	.640
5	.908	.731	.681	.654	.637
7	.868	.728	.679	.652	.636

Table 4.3

$\delta=2$

p \ N/p	1	2	3	4	5
3	1.0	.763	.709	.678	.656
5	.934	.747	.695	.666	.653
7	.977	.739	.689	.666	.652

4.2 The U-method:

Another estimator of ρ_{21} is obtained by the 'leave-one-out method'. Suppose the unknown parameters μ_1, μ_2 and Σ are estimated by $\overset{o}{X}$, $\bar{Y}$ and $S = (m+n-3)^{-1} \overset{\bullet}{A}$ [defined in the proof of Theorem 4.1] and the left out observation X_1, known to belong to $\mathcal{C}_1$, is used for testing. Define

$$(4.22) \qquad V_1^* = \begin{cases} 1 & \text{if } X_1 \text{ is classified as a member of } \mathcal{C}_2 \\ 0 & \text{otherwise} \end{cases}$$

when the linear classifier (2.3) is employed. This process is repeated successively by leaving out $X_2,\ldots,X_m$ and the corresponding $V_2^*,\ldots,V_m^*$ are obtained. Then, an estimate of ρ_{21}, given by the U-method is

$$(4.23) \qquad \varepsilon_4 = \frac{1}{m}\sum_{i=1}^{m} V_i^* \ .$$

Since $X_1, X_2,\ldots,X_m$ are independent and identically distributed, V_i^*'s will also be identically distributed, although not independently. Thus,

$$(4.24) \qquad E(\varepsilon_4) = E(V_1^*)$$

$$= P[\{X_1 - \tfrac{1}{2}(\overset{o}{X}+\bar{Y})\}' \overset{0-1}{S} (X-\bar{Y}) < 0] \ .$$

Lauchenbruch and Mickey (19) have obtained $E(V_1^*)$ by the Monto-Carlo method. An approximate value of this quantity can also be obtained by the following expression of Okamoto (1963).

$$E(V_1^*) \approx \Phi(-\frac{\delta^2}{2}) + \frac{a_1}{m-1} + \frac{a_2}{n} + \frac{a_3}{m+n-3}$$

$$(4.25) \qquad + \frac{b_{11}}{(m-1)^2} + \frac{b_{22}}{n^2} + \frac{b_{12}}{(m-1)n} + \frac{b_{13}}{(m-1)(m+n-3)}$$

$$+ \frac{b_{33}}{(m+n-3)^2}$$

where a_i's and b_i's are certain constants depending upon δ^2 and p, but independent of m and n, δ^2 is the Mahalnobis distance and $\Phi(\cdot)$ represents the distribution fucntion of the standard normal random variable. In table 4.4 below some numerical values are given for selected values of m, n and δ^2. Values given in this table agree with those obtained by Lauchenbruch and Mickey (1968).

The expected values of $E(\varepsilon_3)$ and $E(\varepsilon_4)$ obtained by the above theoretical considerations agree with the well known results obtained from empirical studies that ε_4 is less biased than ε_3 as an estimator of ρ_{21}. Thus the U-method provides a better estimator of ρ_{21} than the C-method [when the criterion of comparison is bias]. From the point of view of calculations the estimator ε_3, given by the C-method, is easier to evaluate than ε_4, although convenience of evaluation is not a meaningful criterion in view of the availability of modern computers.

Table 4.4

$\delta=1$

N/P \ P	1	2	3	4	5
3	0.6526	0.6141	0.6376	0.6507	0.6588
5	0.6505	0.6242	0.6399	0.6506	0.6578
7	0.6713	0.6302	0.6417	0.6510	0.6577

5. VARIANCE OF THE ESTIMATOR OF PROBABILITY OF ERROR BY THE C-METHOD: Σ IS KNOWN.

The estimator of probability of error ρ_{21} on the design set was defined by ε_1 in Section 3 and a method of evaluating its expectation was described there. Although the expectation of ε_3 provides us information regarding the unbiasedness, knowledge of the variance of this estimator will further increase our understanding of the behavior of this estimator. Very little attention has been paid to the variances of estimators of the probability of misclassification.

In this section we consider the variance of ε_1 when Σ is assumed known. Foley (1972) who considered $E(\varepsilon_1)$ for this situation also obtained an approximate expression for the variance of ε. Recall $m\varepsilon_1 = \sum_1^m T_i$ and marginal densities of T_i's are identical. Foley made the assumption that T_i's are approximately independently distributed, thus $m\varepsilon$ is a binomial random variable with variance of ε equal to $E(\varepsilon)(1-E(\varepsilon))m^{-1}$. This approximate variance has an upper bound $(4m)^{-1}$. We evaluate the exact variance of ε_1.

By definition of ε_1, and symmetry in the distributions of T_i's

$$\mathrm{Var}(\varepsilon_1) = \frac{1}{m^2}\{\sum_{1}^{m} \mathrm{Var}(T_i) + \sum_{i\neq j}\sum \mathrm{cov}(T_i, T_j)\}$$

$$= \frac{1}{m^2}\{m\ \mathrm{Var}(T_1) + m(m-1)\ \mathrm{cov}(T_1, T_2)\}$$

(5.1)
$$= \frac{1}{m}\{[ET_1 - E^2T_1] + (m-1)[E(T_1.T_2) - E^2T_1]\}$$

$E(T_1)$ has already been obtained in Section 3. Thus to obtain an exact variance of ε_1, it remains to find $E(T_1.T_2)$ which is obtained in the following subsection.

5.1. Expression of $E(T_1\ T_2)$

We have already made the assumption that Σ is known. Thus, without loss of generality, we will assume that $\Sigma = I$, the identity matrix. On the other hand, because $\Sigma = I$ the linear classifier (2.1) also simplifies and an observation X is classified to C_1 (C_2) if

(5.2) $$\{X - \tfrac{1}{2}(\bar{X}+\bar{Y})\}'\ (\bar{X}-\bar{Y}) > 0\ (<0)\ .$$

In this section we make another simplifying assumption of m = n. It can be seen from the following development that results are easily obtained when $m \neq n$ in exactly a similar manner.

Following the notations of earlier sections, where X_i' denote observations belonging to C_1 and Y_i's belong to C_2, and using the linear classifier (5.2) we obtain,

$$E(T_1\ T_2) = P[X_1 \text{ and } X_2 \text{ are both classified as member of } C_2]$$

$$= P[\{X_i - \tfrac{1}{2}(\bar{X}+\bar{Y})\}'\ (\bar{X}-\bar{Y}) < 0,\ i = 1,2]\ .$$

By a conditional argument, the above probability can also be written as

$$(5.3) \qquad E_{\bar{X},\bar{Y}} \{P[\{X_i - \tfrac{1}{2}(\bar{X}+\bar{Y})\}' (\bar{X}-\bar{Y}) < 0,\ i = 1,2 \mid \bar{X},\bar{Y}]\} \ .$$

First we evaluate the inner term of (5.3) i.e., the conditional probability given $\bar{X}$, $\bar{Y}$. The following lemma proves useful in this evaluation. Let

$$(5.4) \qquad U_i = \{X_i - \tfrac{1}{2}(\bar{X}+\bar{Y})\}' (\bar{X}-\bar{Y}),\ i = 1,2 \ .$$

<u>Lemma 5.1</u>. The joint distribution of (U_1,U_2) given $\bar{X} = \bar{x}$ and $\bar{Y} = \bar{y}$ is a bivariate normal with mean vector $1/2(\alpha,\alpha)$ and the covariance matrix

$$\frac{\alpha}{n} \begin{bmatrix} n-1 & -1 \\ -1 & n-1 \end{bmatrix} ,$$

$$(5.5) \qquad \text{where} \qquad \alpha \equiv \alpha(\bar{x},\bar{y}) = (\bar{x}-\bar{y})' (\bar{x}-\bar{y}) \ .$$

<u>Proof</u>: Let $\hat{X}$ denote the mean of the (n-2) observations $X_3,\ldots,X_n$. Then X_1, X_2 and $\hat{X}$ are statistically independent, normally distributed with the same mean μ and the covariance matrices I, I and $(n-2)^{-1}$ I respectively. To obtain the joint distribution of X_1, X_2 and $\bar{X}$ from the joint distribution of X_1, X_2, $\hat{X}$ we apply the transformation

$$X_1 = X_1$$
$$X_2 = X_2$$
$$\bar{X} = N^{-1}(X_1+X_2+(n-2)\hat{X}) \ .$$

The joint density of X_1, X_2, $\bar{X}$ is given by

$$n^{p/2}(n-2)^{-\frac{p}{2}}(2\Pi)^{-\frac{3p}{2}} \exp -\frac{1}{2}[(x_1-\mu_1)'(x_1-\mu_1) + (x_2-\mu_1)'(x_2-\mu_1)$$

$$+(n-2)\{(n-2)^{-1}(n\bar{x}-x_1-x_2)-\mu_1\}'\{(n-2)^{-1}(n\bar{x}-x_1x_2)-\mu_1\}].$$

Since $\bar{X}$ is a p-variate gaussian random variable, the conditional distribution of (X_1,X_2) given $\bar{X}$ is obtained by dividing the above expression by the pdf of $\bar{X}$. This conditional pdf of X_1,X_2 simplyfies to

$$f(x_1,x_2|\bar{x}) = (2\Pi)^p(\frac{n}{n-2})^{\frac{p}{2}} \exp -\frac{1}{2(n-2)}[(n-1)\{x_1'x_1+x_2'x_2\} + 2x_1'x_2$$

$$+2n\,\bar{x}'\bar{x} - 2n\bar{x}'\ (x_1+x_2)]$$

$$= (2\Pi)^p\ (\frac{n}{n-2})^{\frac{p}{2}} \exp -\frac{1}{2}[\{(x_1-\bar{x})'\vdots(x_2-\bar{x})'\}\Lambda^{-1}\begin{pmatrix} x_1-\bar{x} \\ x_2-\bar{x}\end{pmatrix}]$$

where

$$\Lambda = \frac{1}{n}\begin{bmatrix} (n-1)I & -I \\ -I & (n-1)I \end{bmatrix}.$$

Thus the conditional pdf of $(X_1',X_2')'$ given $\bar{X}$ is a 2p-variate gaussian with mean vector $(\bar{x}',\bar{x}')'$ and covariance Λ. Next, if $Z_i = X_i'\ (\bar{X}-\bar{Y})$; $i = 1,2$, then the conditional density of $(Z_1,Z_2)'$ given $\bar{X}$, $\bar{Y}$ is a bivariate normal with

$$E\left[\begin{pmatrix} Z_1 \\ Z_2\end{pmatrix} \Big|\ \bar{x},\bar{y}\right] = \begin{bmatrix} \bar{x}'(\bar{x}-\bar{y}) \\ \bar{x}'(\bar{x}-y)\end{bmatrix} \text{ and } \operatorname{cov}\left[\begin{pmatrix} Z_1 \\ Z_2\end{pmatrix} \Big|\ \bar{x},\bar{y}\right] = \frac{1}{n}\begin{bmatrix} n-1 & -1 \\ -1 & n-1\end{bmatrix}\ \alpha\ .$$

Finally, if

$$U_i = (X_i - \frac{\bar{X}+\bar{Y}}{2})(\bar{X}-\bar{Y}),\ i = 1,2,$$

then the joint density of $(U_1,U_2)'$, given $\bar{X} = \bar{x}$, $\bar{Y} = \bar{y}$ is also a bivariate normal with mean vector $\frac{1}{2}(\alpha,\alpha)'$ and with the same covariance matrix as of Z's. This completes the proof of the lemma.

From (5.3) $E(V_1V_2) = E\{P(U_1<0, U_2<0|\bar{X},\bar{Y})\}$ where U_i's are defined by (5.4) and using Lemma 4.1 we get

$$P(U_1<0,U_2<0|\bar{X},\bar{Y}) = \int_{-\infty}^{0}\int_{-\infty}^{0} k(\alpha)\exp-\frac{(n-1)}{2(n-2)\alpha}\{(u_1-\frac{\alpha}{2})^2 + (u_2-\frac{\alpha}{2})^2 + \frac{2}{n-1}(u_1-\frac{\alpha}{2})(u_2-\frac{\alpha}{2})\}du_1\, du_2$$

$$= \int_{\alpha}^{\infty}\int_{\alpha}^{\infty} k(\alpha)\exp-\frac{(n-1)}{2(n-2)\alpha}\{t_1^2 + t_2^2 + \frac{2}{n-1}t_1t_2\}\, dt_1\, dt_2$$

$$= 2k(\alpha)\int_0^{\frac{\pi}{4}}\int_{\frac{\alpha}{2\sin\theta}}^{\infty}[\exp-\frac{(n-1)r^2}{2(n-2)\alpha}\{1+\frac{\sin 2\theta}{n-1}\}]r\, dr\, d\theta$$

$$= 2k(\alpha)\int_0^{\frac{\pi}{4}}\frac{(n-2)\alpha}{(n-1)\{1+\frac{\sin 2\theta}{n-1}\}}\exp\{-\frac{1}{8}\frac{n-1}{n-2}(1+\frac{\sin 2\theta}{n-1})\frac{\alpha}{\sin^2\theta}\}\, d\theta$$

$$(5.6)\qquad = \frac{\sqrt{n(n-2)}}{\pi}\int_0^{\frac{\pi}{4}}\{\exp-\frac{1}{8}\frac{n-1}{n-2}(1+\frac{\sin 2\theta}{n-1})\frac{\alpha}{\sin^2\theta}\}\frac{d\theta}{(n-1+\sin 2\theta)}$$

where

$$k(\alpha) = \sqrt{n}\,(2\pi\alpha\sqrt{n-2})^{-1}.$$

In the second expression above we use the symmetry of the integrand to change the integral from the third quadrant to the first quadrant. To get the third equality above, first we use the symmetry of the integrand around the line $t_1 = t_2$ and then change to polar coordinate system.

Thus to evaluate $E(T_1\ T_2)$ it remains to take the expectation of (5.6) with respect to $\bar{X}$, $\bar{Y}$. Since (5.6) depends on $\bar{X}$, $\bar{Y}$ only through α, we take the expectation with respect to α. But $(n/2)\alpha$ is a noncentral chisquare random variable with p degrees of freedom and with noncentrality parameter $\frac{n}{2}\delta^2$ where $\delta^2 = (\mu_1-\mu_2)'\ (\mu_1-\mu_2)$. Therefore,

$$E(V_1V_2) = \frac{n}{2}\int_0^\infty \Big\{ \sum_{s=0}^{\infty} \exp(-\frac{n}{4}\delta^2)\ \frac{1}{s!}\ (\frac{n}{4}\delta^2)^s\ 2^{-(s+\frac{p}{2})}\ \Gamma^{-1}(s+\frac{p}{2})$$

$$(\exp-\frac{n\alpha}{4})(\frac{n\alpha}{2})^{s+\frac{p}{2}-1} \Big\}\Big\{\sqrt{\frac{n(n-2)}{\pi}} \int_0^{\frac{\pi}{4}} [\exp-\frac{1}{8}\ \frac{(n-1+\sin 2\ )}{(n-2)\sin^2\theta}\ \alpha]$$

$$\frac{d\theta}{(n-1+\sin 2\theta)}\Big\}\ d\alpha.$$

Due to convergence of the above integral, the order of integration can be interchanged. Interchanging the order of integration and then integrating over α produces

$$E(V_1V_2) = \int_0^{\frac{\pi}{4}} \{ \sum_{s=0}^{\infty} \exp(-\frac{n}{4}\delta^2)\frac{1}{s!}(\frac{n}{4}\delta^2)^s \, 2^{-(s+\frac{p}{2})} \, \Gamma^{-1}(s+\frac{p}{2}) \frac{\sqrt{n(n-2)}}{\pi}$$

$$\Gamma(s+\frac{p}{2}) \, [1+\frac{1}{2n}\{\frac{n-1+\sin 2\theta}{(n-2\sin^2\theta}\}^{-(s+\frac{p}{2})} \} \frac{d\theta}{(n-1)+\sin 2\theta}$$

$$(5.7) \qquad = \frac{\sqrt{n(n-2)}}{\pi} \int_0^{\frac{\pi}{4}} \exp-\frac{n\delta^2}{4}\{\frac{n-1+\sin 2\theta}{2n(n-2)\sin^2\theta+(n-1)+\sin 2\theta}\}$$

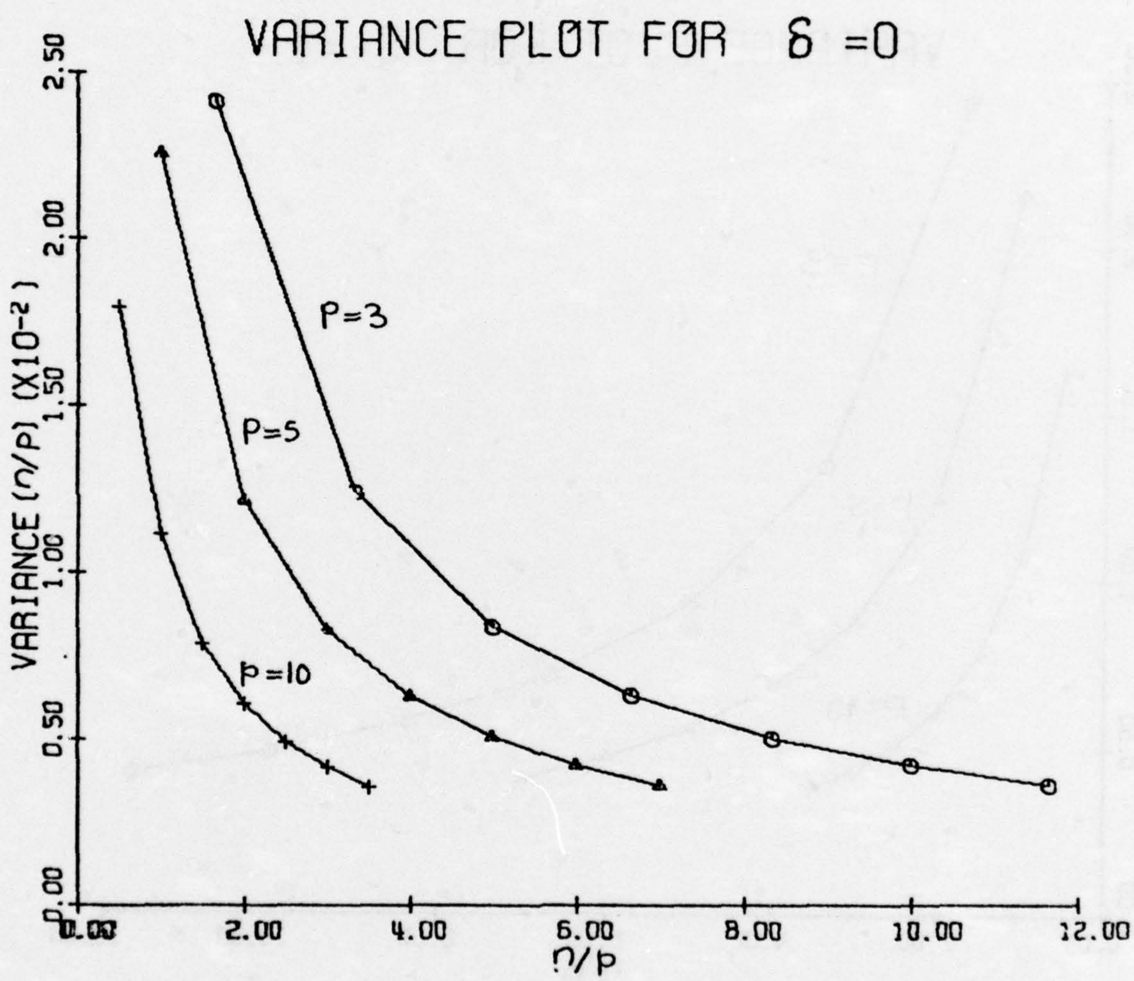
$$\{1+\frac{1}{2n}\frac{n-1+\sin 2\theta}{(n-2)\sin^2\theta}\}^{-\frac{p}{2}} \frac{d\theta}{(n-1)+\sin 2\theta} \; .$$

Expression (5.7) involves integration over one variable and to evaluate it we use numerical integration. The graph of var(ε_1) is also presented for various values of N/p.

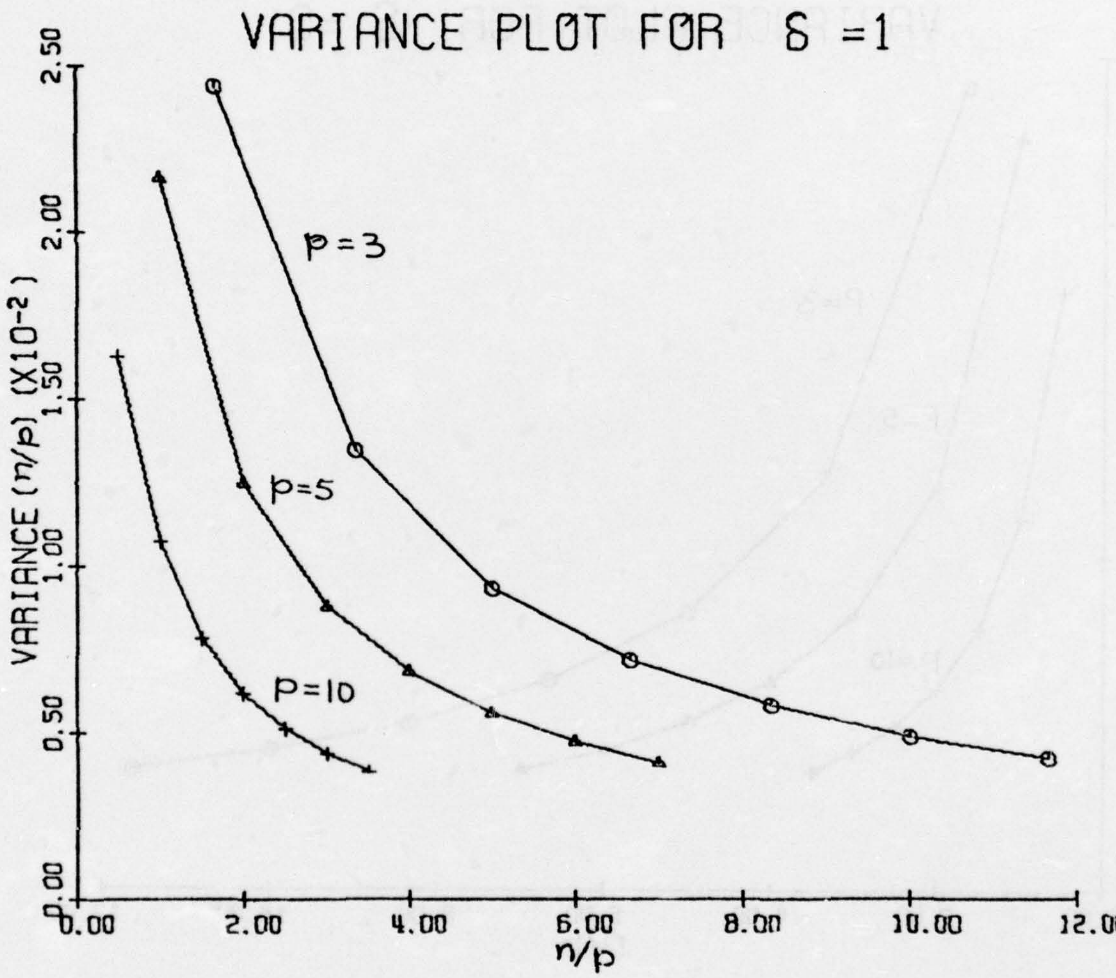
5.2. Some Numerical Values of Var(ε)

As shown earlier [see expression (5.1)] to evaluate var(ε_1) we needed to know E(T_1) and E(T_1 T_2). To evaluate ET_1 we use equation (3.5) and E(T_1T_2) can be obtained by numerical integration of (5.7). Thus var(ε_1) can be calculated exactly. In Table 5.1 we present these values for some choices of n and p. The values in Table 5.1 show an interesting feature, namely that for fixed value of n if p increases then var(ε_1) decreases. Thus, for fixed n although the bias in ε_1 increases, the variance of ε_1 decreases as p increases.

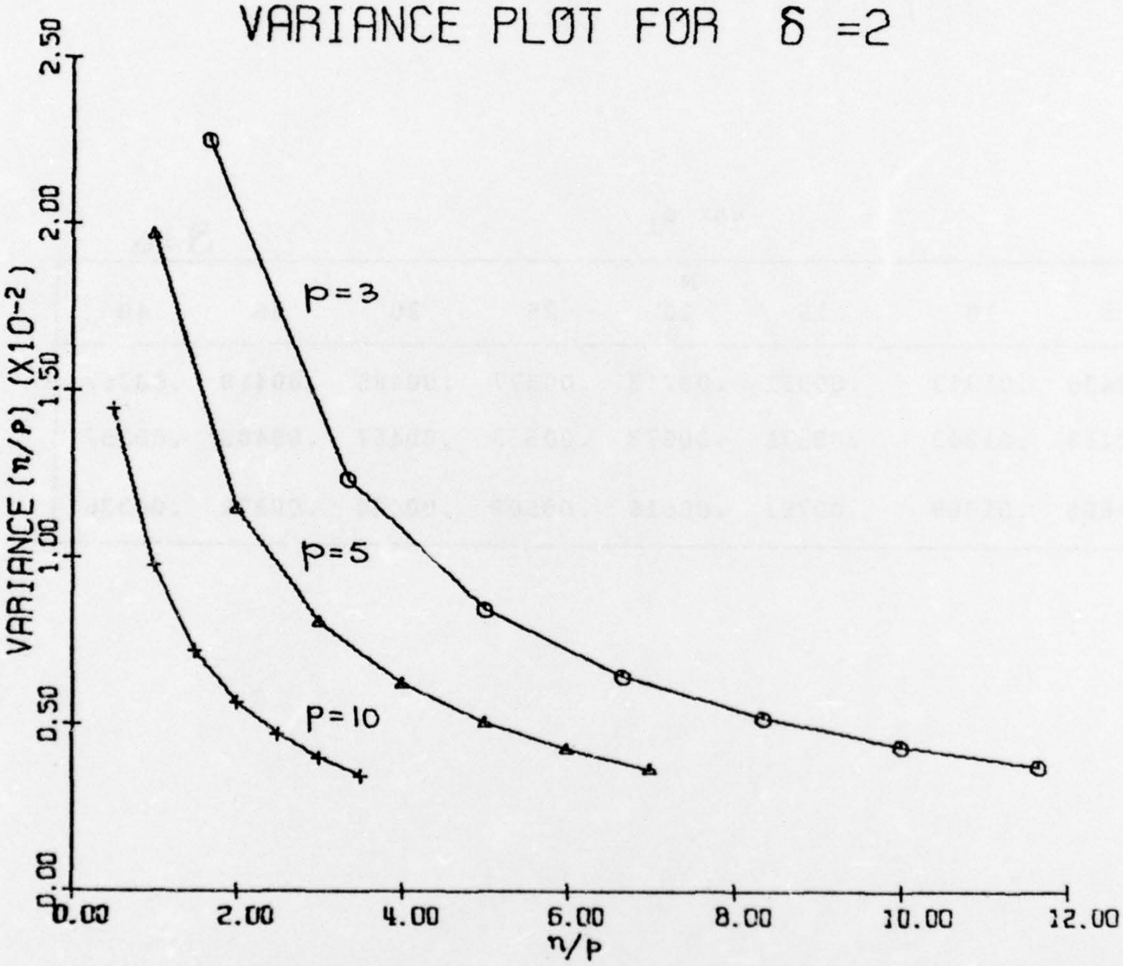
A computer program which evaluates the var(ε_1) is given in the Appendix.



VARIANCE PLOT FOR δ =0
VARIANCE (n/P) (X10⁻²)
n/P
P=3
P=5
P=10
0.00
0.50
1.00
1.50
2.00
2.50
2.00
4.00
6.00
8.00
10.00
12.00



VARIANCE PLOT FOR δ =1
VARIANCE (n/p) (X10-2)
n/p
p=3
p=5
P=10
0.00
0.50
1.00
1.50
2.00
2.50
0.00
2.00
4.00
6.00
8.00
10.00
12.00


VARIANCE PLOT FOR δ =2
VARIANCE (n/p) (X10-2)
2.50
2.00
1.50
1.00
0.50
0.00
0.00
2.00
4.00
6.00
8.00
10.00
12.00
n/p
p=3
p=5
p=10

var e_1

$\delta = 0$

p	N 5	10	15	20	25	30	35	40
3	.02436	.01343	.00931	.00713	.00577	.00485	.00418	.00367
5	.02158	.01243	.00876	.00678	.00553	.00467	.00405	.00357
10	.01625	.01069	.00781	.00616	.00509	.00434	.00379	.00336

6. VARIANCE OF THE ESTIMATE OF PROBABILITY OF ERROR BASED ON THE U-METHOD.

In this section an approximate value of variance of the estimate of the probability of error of misclassification is obtained when $\sum$ is assumed known. This estimator ε_2 was defined in (3.6).

$$\text{(6.1)}\quad \operatorname{Var}(\varepsilon_2) = \frac{1}{m}[\{ET_1^* - E^2(T_1^*)\} + (m-1)\{E(T_1^*T_2^*) - E^2(T_1^*)\}] \ .$$

$E(T_1^*)$ has already been evaluated [see equation (3.8)]. Thus, once again, to calculate the $\operatorname{var}(\varepsilon_2)$ it remains to evaluate $E(T_1^*T_2^*)$.

6.1. Expression for $E(T_1^* \ T_2^*)$

By definition

$$\text{(6.2)}\qquad T_1^* = \begin{cases} 1 & \text{if } \{X_1 - \frac{1}{2}(\overset{o}{\bar{X}}+\bar{Y})\}'(\overset{o}{\bar{X}}-\bar{Y}) < 0 \\ 0 & \text{otherwise} \end{cases}$$

and a similar expression holds for T_2^*. At this stage let

$$\bar{X}^{(i)} = (m-1)^{-1}[m\bar{X}-X_i] = (m-1)^{-1}\sum_{j\neq i} X_j \quad .$$

Then

$$E(T_1^*T_2^*) = P[\{X_i - \frac{1}{2}(\bar{X}^{(i)}+\bar{Y})\}' \ (\bar{X}^{(i)}-\bar{Y}) < 0; \ i = 1,2]$$

$$\text{(6.3)}\qquad = P[\{X_i - \frac{1}{2}[(m-1)^{-1}(m\bar{X}-X_i)+\bar{Y}]\}'\{(m-1)^{-1}(m\bar{X}-X_i)-\bar{Y}\} < 0, \ i = 1,2] \ .$$

By a sequence of arguments presented below the events of interest in (6.3) can be written in a convenient form. Note that for $i = 1,2$,

$$\{X_i - \tfrac{1}{2}[(m-1)^{-1}(m\bar{X}-X_i)+\bar{Y})\}'\{(m-1)^{-1}(m\bar{X}-X_i)-\bar{Y}\} < 0$$

$$\text{iff } [\frac{2m-1}{2(m-1)} X_i - \frac{1}{2}(\frac{m}{m-1}\bar{X} + \bar{Y})]' \; [-\frac{1}{m-1} X_i + (\frac{m}{m-1}\bar{X}-\bar{Y})] < 0$$

$$\text{iff } [-\frac{2m-1}{2(m-1)^2} X_i'X_i + X_i'\{(\frac{m}{m-1})^2\bar{X} - \bar{Y}\} - \frac{1}{2}\{(\frac{m}{m-1})^2\bar{X}'\bar{X} - \bar{Y}'\bar{Y}\} < 0$$

$$\text{iff } [X_i'X_i - \frac{2(m-1)^2}{2m-1}\{(\frac{m}{m-1})^2\bar{X}-\bar{Y}\}' X_i + \frac{(m-1)^2}{2m-1}\{(\frac{m}{m-1})^2\bar{X}'\bar{X} - \bar{Y}'\bar{Y}\} > 0]$$

$$\text{(6.4)} \qquad \text{iff } [Z_i'Z_i > \frac{m^2(m-1)^2}{(2m-1)^2}\{\bar{X}-\bar{Y}\}'\{\bar{X}-\bar{Y}\}$$

where $Z_i = X_i - \frac{(m-1)^2}{(2m-1)}\{(\frac{m}{m-1})^2\bar{X} - \bar{Y}\}$. Therefore by (6.3) and (6.4) *and using a conditional argument, we obtain*

$$\text{(6.5)} \qquad E(T_1^*,T_2^*) = E[P\{Z_i'Z_i > \{\frac{m(m-1)}{2m-1}\}^2 (\bar{X}-\bar{Y})'(\bar{X}-\bar{Y});\ i = 1,2|\bar{X},\bar{Y}\}]$$

In the proof of Lemma 5.1 we have shown that the joint distribution of $(X_1',X_2')'$ given $\bar{X}$ is a 2p-dimensional gaussian distribution with mean vector $(\bar{X}',\bar{X}')'$ and covariance matrix Λ. Therefore, conditional distribution, given $\bar{X}$, $\bar{Y}$, of $(Z_1',Z_2')'$ is also a 2p-dimensional gaussian with mean vector

$$\frac{(m-1)^2}{(2m-1)}\begin{bmatrix} (\bar{Y} - \bar{X}) \\ \\ (\bar{Y} - \bar{X}) \end{bmatrix}$$

and covariance matrix Λ. From the above development, it is obvious that Z_1 and Z_2 are dependent random variable and the marginal distributions of $\frac{m}{m-1} Z_i'Z_i$ are noncentral chisquares each with p degrees of freedom and each with noncentrality parameters $m(m-1)^3 \ (2m-1)^{-2} (\bar{X}-\bar{Y})' (\bar{X}-\bar{Y})$; given $\bar{X}$ and $\bar{Y}$.

If we make the approximation of treating Z's as independent random variables, then to this degree of approximation, [which is expected to be small when m is large, because covariance between Z's is $-(m-1)^{-1}I$],

$$(6.6) \qquad E(T_1^*T_2^*) \approx E[P^2\{Z_1'Z_1 > (\frac{m(m-1}{2m-1})^2 \ \alpha(\bar{X},\bar{Y}) | \alpha(\bar{X},\bar{Y})\}]$$

where $\alpha(\bar{X},\bar{Y})$ is defined in (5.5). As seen in the last section, $\frac{m}{2}\alpha(\bar{X},\bar{Y})$ itself follows a chisquare distribution, with p degrees of freedom and noncentrality parameter $\frac{m}{2}(\mu_1-\mu_2)' \ (\mu_1-\mu_2) = \delta^2$. Using these distributional properties, further simplification of (6.6) is presented below.

Set $c = \frac{m^3(m-1)}{(2m-1)^2}$ and $d = \frac{m(m-1)^3}{(2m-1)^2}$. Then,

$$P[Z_1'Z_1 > (\frac{m(m-1)}{2m-1})^2 \alpha(\bar{X},\bar{Y}) | \alpha(\bar{X},\bar{Y})]$$

$$= \sum_{j=0}^{\infty} e^{-\frac{d\alpha}{2}} (\frac{d\alpha}{2})^j \frac{1}{j!} \int_{c\alpha}^{\infty} \frac{e^{-\frac{x}{2}} x^{\frac{p+2j}{2}-1}}{\Gamma(\frac{p+2j}{2}) 2^{\frac{p+2j}{2}}} dx$$

$$\therefore g(\alpha) = P^2[Z_1'Z_1 > (\frac{m(m-1)}{2m-1})^2 \alpha(\bar{X},\bar{Y}) | \alpha(\bar{X},\bar{Y})]$$

$$= \sum_{j,k=0}^{\infty} e^{-d\alpha} (\frac{d\alpha}{2})^{j+k} \frac{1}{j!k!} \int_{c\alpha}^{\infty} \int_{c\alpha}^{\infty} \frac{e^{-\frac{1}{2}(x+y)} x^{\frac{p+2j}{2}-1} y^{\frac{p+2k}{2}-1}}{\Gamma(\frac{p+2j}{2}) \Gamma(\frac{p+2k}{2}) 2^{p+j+k}} dxdy.$$

By changing to the polar coordinates, i.e.,

$$x = r\cos\theta$$

$$y = r\sin\theta$$

we obtain

$$\vartheta(\alpha) = \sum_{j,k=0} C\, e^{-d\alpha} \alpha^{j+k} \int_0^{\frac{\pi}{4}} \int_{c\alpha/\cos\theta}^{\infty} [\{\sin\theta^{\frac{2k+p}{2}-1} \cos\theta^{\frac{2j+p}{2}-1} +$$

$$+ \cos\theta^{\frac{2k+p}{2}-1} \sin\theta^{\frac{2j+p}{2}-1}\}\, e^{-\frac{r}{2}(\cos\theta+\sin\theta)}\, r^{j+k+p-1}\, dr\, d\theta]$$

$$= \sum_{j,k=0}^{\infty} C\, e^{-d\alpha} \alpha^{j+k} \int_0^{\frac{\pi}{4}} \sin\theta^{\frac{2k+p}{2}-1} \cos\theta^{\frac{2j+p}{2}-1} + \cos\theta^{\frac{2k+p}{2}-1} \sin\theta^{\frac{2j+p}{2}-1}$$

$$\{(\frac{2}{\sin\theta+\cos\theta})^{j+k+p} \sum_{i=0}^{j+k+p-1} \frac{\Gamma(j+k+p)}{\Gamma(i+1)}\, e^{-\frac{c\alpha}{2}(1+\tan\theta)} [\frac{c\alpha}{2}(1+\tan\theta)]^{i}\}d\theta$$

where $C = (\frac{d}{2})^{j+k} \dfrac{1}{j!k!\,\Gamma(\frac{p+2j}{2})\,\Gamma(\frac{p+2k}{2})\,2^{p+j+k}}$.

Finally,

$$E(T_1^* T_2^*) = E(\vartheta(\alpha))$$

$$= \sum_{\ell=0}^{\infty} \sum_{j,k=0}^{\infty} \sum_{i=0}^{j+k+p-1} C^* \int_0^{\infty} \int_0^{\frac{\pi}{4}} \{e^{-\frac{m}{4}\alpha} (\frac{m\alpha}{4})^{\frac{2\ell+p}{2}-1} e^{-\frac{c\alpha}{2}(1+\tan\theta)}$$

$$e^{-d\alpha}\, \alpha^{j+k+i}\, q(\theta)]d\theta\, d(\frac{m}{4}\alpha)$$

where $C^* = C2^{j+k+p} \dfrac{\Gamma(j+k+p)}{\Gamma(i+1)} e^{-\frac{\delta^2}{2}} (\frac{\delta^2}{2})^{\ell} \frac{1}{\ell!} (\frac{c}{2})^{i} \dfrac{1}{\Gamma(\frac{2\ell+p}{2})}$

and

$$q(\theta) = \{\sin\theta^{\frac{2k+p}{2}-1} \cos\theta^{\frac{2j+p}{2}-1} + \cos\theta^{\frac{2k+p}{2}-1} \mathrm{sim}\theta^{\frac{2j+p}{2}-1}\}$$

$$\{\sin\theta+\cos\theta\}^{-(j+k+p)} \{1+\tan\theta\}^{i} .$$

Interchanging the order and then integrating over α, in the above expression produces,

$$(6.7) \qquad E(V_1^* V_2^*) = \sum_{\ell,j,k=0}^{\infty} \sum_{i=0}^{j+k+p-1} C^{**} \int_0^{\frac{\pi}{4}} q^*(\theta)\,d\theta$$

where

$$C^{**} = C^* \left(\frac{m}{4}\right)^{\ell+\frac{p}{2}} \Gamma\left(\ell+j+k+i+\frac{p}{2}\right)$$

and

$$q^*(\theta) = q(\theta)\left[\frac{m}{4}+d+\frac{c}{2}(1+\tan\theta)\right]^{-\left(\frac{p}{2}+\ell+j+k+i\right)} .$$

This last integral can be evaluated numerically.

At the present time we have not been able to calculate the value of $E(T_1^* T_2^*)$ even with the help of a computer. Therefore, the numberical values and a comparison of the two estimates of error of probability will be presented in a forthcoming technical report.

APPENDIX

The following computer program provides the expectation and the variance of the estimate of the error of probability on the design set. A plotting algorithm is also attached.

MAIN

```
COMPUTE THE MEAN AND VARIANCE OF THE PROBA. OF MISCLASSIFICATION
EXTERNAL  F2
COMMON/CONS1/C,D,AN,IP,I,J,K,L,JK,HP1,D1,DD
COMMON/GAMA/GI1,GJ1,GK1,GL1,GJHP,GKHP,GJKP,GIJKL
DATA      PI4,AERR,RERR,EPS/.7853981634E0,1.0E-6,1.0E-4,1.0E-1/
DATA  ZERO,ONE,TWO/0.0E0,1.0E0,2.0E0/
DO 600  ID-1,1
IDS=ID-1
WRITE(6,920)
DO 500  I1=1,1
READ(5,900) IP
HP=FLOAT(IP)/TWO
HP1=HP-ONE
DO 480  12=1,1
IN=12*5
AN=FLOAT(IN)/TWO/TWO
D1=FLOAT(IDS*AN
AN1=TWO*IN-ONE
AN2=FLOAT(IN*(IN-1))
AN3=D1*AN2/TWO
AN4=ONE/SQRT(AN1)
AN5=ONE/SQRT(AN1+TWO*TWO*AN2)
RHO=-AN4*AN5
AC=ONE-RHO)/TWO
AL1=AN3*(AN4-AN5)**2/(ONE+RHO)
AL2=AN3*(AN4+AN5)**2/(ONE-RHO)
T1-EXP(-AL1-AL2)
T2=ONE
EV1=ZERO
DO 300  IP=1,21
AA=HP1+IR-ONE
T3=ONE
EVO=ZERO
DO 200  IS=1,21
AB=HP1+IS-ONE
CALL  MDBETA(AC,AA,AB,T4.IER)
EV=(ONE-T4)*T2*T3
```

```
EVO=EVO+EV
IF(EV.LE.1.OE-1o) GO TO 210
T3=T3*AL2/IS
CONTINUE
T2=T2*AL1/IR
EV1=EV1+EVO
IF(EVO.LE.1.OE-6) GO TO 310
CONTINUE
EV1=EV1*T1
R=AN2/AN1**2
C=B*IN*IN?TWO
D=b89in-1)*(IN-1)
CC=C/TWO
CSS=FXP(-D1)*AN**HP
TLL*9
IF(IDS.EQ.0)  ILL=1
EV1V2=ZERO
DO 450  IL=1,I11
L=IL-1
GL1=GAMMA(FLOAT(IL))
FVV1=ZERO
DO 440  IJ=1,11
J=IJ-1
GJ1=GAMMA(FLOAT(IJ))
GJHP=GAMMA(J+HP)
EVVO=ZERO
DO 420 IK=1,11
K=IK-1
GK1=GAMMA(FLOAT(IK))
GKHP=GAMMA(K+HP)
JK=J+K
JKPJK+IP
GJKP=GAMMA(FLOAT(JKP))
EVV=ZERO
DO 400  II=1,JKP
I=II-1
GI1=GAMMA(FLOAT(II))
GIJKL=GAMMA(I+JK+L+HP)
ETEMP=CSS*DCADRE(F2,EPS,PI4,AERR,RERR,ERROR,IER)
EVV=EVV+ETEMP
IF(ETEMP.LE.AERR) GO TO 410
CONTINUE
EVVO=EVVO+EVV
IF(EVV.LE.8.OE-6) GO TO 430
CONTINUE
EVV1=EVV1+EVVO
IF(EVVO.LE.5.OE-5) GC TO 445
CONTINUE
```

MAIN

```
VARE=(EV1-EV1*EV1+(IN-1)*(EV1V2-EV1*EV1))/IN
WPITE(6,950) EV1,EV1V2,VARE,IP,IN,IDS
CONTINUE
CONTINUE
FORMAT(3I5)
FORMAT('  IP=',2(I3,2X),4(E14.7,2X))
FORMAT(7X,'EV1',11X,'EV1V2',12X'VARE',8X,'IP',3X,'IN',2X,'IDS')
FORMAT(2X,'EVV1=',2(E15.3,2X),3(I5,2X))
FORMAT(2X,3(E14.7,2X),3(I3,2X))
STOP
END
```

```
          COMPUTE THE EXPECTED VALUE AND VARIANCE OF THE PROBA. OF
          ERROR ON THE DESIGN SET
          E(V1) =SUM(EXP(-LAMDA**2/2.)*(1./FACT R))*(LAMDA**2/2.)**
          *GAMMA(R+L/2.+)0.50)*I(L+2R,N)/GAMMA(0.50)/GAMMA L/2
          SUM OVER R=0 TO INFINITY
          INPUT IL,IN,ID1
      DCADRE IS AN IMSL LIBRARY FUNCTION WHICH INTEGRATE E, K)
      USING CAUTIOUS ADAPTIVE ROMBERG EXTRAPOLATION
      MDBETA IS AN IMSL LIBRARY SUBROUTINE WHICH DOES INCOMPLETE B
      PROBABILITY DISTRIBUTION FUNCTION INTEGRATION
```

```
              EXTERNAL  F1
              COMMON  /INPUT1/IL,IN,ID1
              DATA    A,AERR,RERR,PI4/1.0E-4,1.0E-6,1.0E-4,.7853981634E0/
              DATA  IPEN,IFLAG,ICOMNT,NI/1,0,1,7/
              DATA HALF,ONE,TWO/.5E0,1.0E0,2.0E0/
              REAL  XLABL 19),YLABL 19),TITL(19),ABSC(7),ORD V)
              REAL  X(4)/0.0.12.0,6.0,0.0/.Y 4)/0.0,0.025.5.0,0.0/
              CALL PLOTID
              CALL PLOT 0.0,2.5,-3)
              READ(5,900) NXL,XLABL,NYL,YLABL
              DO 600  ID=1,3
              ID1=ID-1
              WRITE (6,920)
              READ (5,900) NTL,TITL
              CALL  GRAPHS(X,Y,XLABL,NXL,YLABL,NYL,TITL,NTL)
              DO 500  I1=1,3
              ISYMB=I1
              READ (6,910) IL
              DO 400  I2=1,7
              IN=I2*5
              D1=FLOAT(IN*ID1)/TWO/TWO
              T1=EXP(-D1)/TWO
              AN1=(TWO*IN-TWO)/(TWO*IN-ONE)
              IR=0
              EVO=0.0E0
              T2=ONE
        150   S=FLOAT(IL)/TWO+FLOAT(IR)
              CALL  MDBET AN1,S,HALF,T3,IER)
              EVR=T2*T3
              EVO=EVO+EVR
              IF EVR.LE.AERR) GO TO 200
              IR=IR+1
              T2=T2*D1/FLOAT(IR)
              GO TO 150
        200   EV1=EVO*T1
              EV1V2=DCADRE(F1,API4,AERR,RERR,ERROR,IER)
              VARE=(EV1-EV1*EV1+(IN-1)*(EV1V2-EV1*EV1))/IN
              ORD(I2)=VARE
              ABSC(I2)=FLOAT(IN)/FLOAT IL)
        400   WRITE(6,940) IL,IN,ID1,ABSC(I2),ORD(I2),EV1,EV1V2
              CALL DATA$(ABSC,ORD,NI,IPEN,IS7MB,IFLAG,ICOMNT)
              CALL GRAPH$(X,Y,XLABL,NXL,YLABL,NYL,TITL,NTL)
FORTRAN IV G LEVEL  21                MAIN              DATE = 76014
```

```
  500 CALL DATA$(ABSC,ORD,NI,IPEN,ISYMB,IFLAC,ICOMNT)
  600 CALL PLOT (10.0,0.0,-3)
      CALL FRAME$
  900 FORMAT(I4,19A4)
  910 FORMAT (S)
  920 FORMAT(/,2X,'IL',3X,'IN',3X,'ID1',8X,'N/L',13X,'VARE',12X,
     *          'EV1',12X,'EV1V2',/)
  940 FORMAT (X,3(I3,2X),4 E14.7,2X))
 1000 STOP
```

blank

A-6

References

1. Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis, Wiley, New York.

2. Das Gupta, S. (1974). "Probability inequalities and errors in misclassification" Annals of Statist. 2, 751.

3. Foley, D. H. (1972). "Considerations of sample and feature size" IEEE Trans. on Information Theory 18, 618.

4. Fukunaga, K. and Kessel, D. L. (1971). "Estimation of classification errors" IEEE Trans. on Computers.

5. Fukunaga, K. (1972). Introduction to Statistical Pattern Recognition. Academic.

6. Mighleyman, W. M. (1962). "The design and analysis of pattern recognition experiments". Bell Sys. Tech. J. 41, 723.

7. Mills, M. (1966). "Allocation rules and their error rates". J. Royal. Statist. Soc. Ser. B. 28, 1.

8. John, S. (1961). "Errors in discrimination". Annals. Math. Statist. 32.

9. Lachenbruch, P. A. (1965). "Estimation of error rates in discriminant analysis" Ph.D. dissertation Univ. of California, Los Angeles.

10. Lachenbruch, P. A. and Mickey, M. R. (1968). "Estimation of error rates in discriminant analysis". Technometrics, 10.

11. Moore, D. S., Whitsitt, S. J. and Landgrebe, D. A. (1976). "Variance comparisons for unbiased estimators of probability of correct classification". IEEE Trans. on Information Theory 22, 102.

12. Okomoto, M. (1963). "An asymptotic expansion for the distribution of the linear discriminant function". Annals. Math. Statist. 34, 1286. [All see Okomoto (1968) A.M.S. 39, 135].

13. Pearson, E. S. and Hartley, H. O. Biometrika Tables for Statisticians, Vol. 1. Cambridge University Press, 1956.

14. Sorem, M. J. (1971). "Estimating the conditional probability of misclassification". Technometrics 13, 333.

15. Sorem, M. J. (1972). "Three probabilities of misclassifications". Technometrics 14, 309.